

# METHOD AND APPARATUS FOR OPTIMIZING A JPEG IMAGE USING REGIONALLY VARIABLE COMPRESSION LEVELS

## FIELD

5 The present invention relates to a method and apparatus  
for optimizing a JPEG image using regionally variable compression  
levels.

## BACKGROUND

10 Compression is a useful method for reducing bandwidth  
consumption and download times of images sent over data networks.  
A variety of algorithms and techniques exist for compressing  
images. JPEG, a popular compression standard that is  
15 particularly good at compressing photo-realistic images, is in  
common use on the Internet. This standard, defined in "JPEG  
Still Image Data Compression Standard", by W.B. Pennebaker and  
J.L. Mitchell, Chapman & Hall, 1992, is based on a frequency  
domain transform of blocks of image coefficients. As seen in  
Figure 1, JPEG calls for subdividing an image frame 12 into 8x8  
pixel blocks 11 and at box 16 transforming the array of pixel  
20 values in each block 11 with a discrete cosine transform (DCT) so  
as to generate 64 DCT coefficients corresponding to each pixel  
block 11. The coefficients for each block 11 are quantized in  
quantization block 20 using a 64 element quantization table 24.  
Each element of table 24 is an integer value from 1 to 255, which  
25 specifies the step size of the quantizer for the corresponding  
DCT coefficients. The quantized coefficients for each block are  
entropy encoded in entropy coding box 28, which performs a

lossless compression. The entropy encoder **28** is coupled to the output of the quantizer **20** from which the former receives quantized image data. Standard JPEG entropy coding uses either Huffman coding or arithmetic coding using either predefined  
5 tables or tables that are computed for a specific image.

The JPEG compressed image data is decompressed by the bottom circuit of Figure **1** by being first passed through an entropy decoder **30**. Next inverse quantization in block **32** using quantization table **34** is performed. Finally the inverse DCT transform block **36** performs an inverse DCT operation to produce the image pixel intensity data.

More specifically, the discrete cosine transform block uses the forward discrete cosine function (DCT) to transform the image pixel intensity  $A(x,y)$  to DCT coefficients  $Y_{mn}$  as follows:

$$Y_{mn} = 1/4 C(m)C(n) \left[ \sum_{x=0}^7 \sum_{y=0}^7 A(x,y) \cos \frac{(2x+1)m\pi}{16} \cos \frac{(2y+1)n\pi}{16} \right]$$

where  $C(m)$  and  $C(n) = 1/\sqrt{2}$  for  $m, n=0$ , and  $C(m)$  and  $C(n) = 1$  otherwise.

The next step is to quantize the DCT coefficients using a quantization matrix, which is an 8 x 8 matrix of step sizes with one element for each DCT coefficient. A tradeoff exists

between the level of image distortion and the amount of compression, which results from the quantization. A large quantization step produces large image distortion, but increases the amount of compression. A small quantization step produces  
5 lower image distortion, but results in a decrease in the amount of compression. JPEG typically uses a much higher step size for the coefficients, which correspond to high spatial frequency in the image, with little noticeable deterioration in the image quality because of the human visual system's natural high frequency rolloff. The quantization is actually performed by dividing the DCT coefficient  $Y_{mn}$  by the corresponding quantization table entry  $Q_{mn}$  and the result rounded off to the nearest integer according to the following:

$$T_{mn} = \text{round}(Y_{mn}/Q_{mn})$$

to give a quantized coefficient  $T_{mn}$ . This type of quantizer is sometimes referred to as a midtread quantizer. An approximate reconstruction of  $Y_{mn}$  is effected in the decoder by multiplying  $T_{mn}$   
20 by  $Q_{mn}$  to obtain a reconstructed  $Y_{mn}$ . The difference between  $Y_{mn}$  and  $Y_{mn}$  represents lost image information causing distortion to be introduced. The amount of this lost information depends on the magnitude of  $Q_{mn}$ .

25 In the case of an image with multiple color channels, the aforementioned steps are applied in a similar fashion to each channel independently. In some cases, some of the color channels

may be sub-sampled to achieve greater compression, without significantly altering the quality of the image reconstruction.

The quantization step is of particular interest since this is where information is discarded from the image. Ideally, one would like to discard as much information as possible, thereby reducing the stored image size, while at the same time maintaining or increasing the image fidelity. Within the standard there is no prescribed method of quantizing the image, but there is nonetheless a popular approach used in the software of the Independent JPEG Group (ISO/IEC JTC1 SC29 Working Group 1), and employed extensively by the general community. This method involves scaling a predetermined quantization table (calculated from statistical importance of basis vectors over a large set of images) by a factor dependent on a user-set quality, which lies in the range 1-100. This method yields good results on average, but is based on statistical averages over many images, and doesn't address global image characteristics, let alone local characteristics.

V. Ratnakar and M. Livny. "RD-OPT: An efficient algorithm for optimizing DCT quantization tables." Proceedings DCC'95 (IEEE Data Compression Conference), pages 332-341, 1995 (and also U.S. Patent No. 5,724,453) describe a rate-distortion dynamic programming optimization technique to reduce distortion for a given target bit-rate, or reduce bit-rate for a given target distortion. This reference uses "Mean Squared Error" as a

measure of distortion and introduces some novel techniques for  
estimating bit-rate that improve the computational efficiency of  
the calculation. This algorithm is designed to calculate a  
single quantization table Q for each channel of the image, and it  
5 is based solely on global aggregate statistics. Also it does not  
take into account varying local image statistics. Moreover the  
method is computationally expensive. There exists another  
technique, which simultaneously optimizes the quantization and  
entropy encoding steps yielding a completely optimum JPEG file  
10 stream. This technique, however is extremely slow and  
unrealistic for real-time JPEG optimization.

US Patent No. 5,426,512 entitled "Image data  
compression having minimum perceptual error" uses a rate-  
15 distortion dynamic programming optimization technique to reduce  
distortion for a target bit-rate, or reduce bit-rate for a target  
distortion. This technique is very similar in concept to V.  
Ratnaker et al., except that the latter uses a "perceptual error"  
measure which attempts to mimic the eye's sensitivity to error.  
20 This algorithm is designed to calculate a single quantization  
table Q for each plane of the image, and it is based solely on  
global aggregate statistics, and it does not take into account  
varying local image characteristics.

25 US Patent No. 5,883,979 entitled "Method for selecting  
JPEG quantization tables for low bandwidth applications" is  
directed mainly at preserving text features in JPEG images at

very low bit-rates. It uses image analysis based on global statistics to determine which DCT basis vectors are more visually important to the image, and weights them accordingly in the quantization table. Again, this algorithm is based on global statistics and also it is geared specifically for preserving textual data in JPEG images.

Ideally, one would like to have an optimal quantization table for every significantly different region of the image (a technique adopted for example in MPEG), which would then allow one to increase image fidelity as a function of file size; this technique of using different quantization tables for different areas of an image is generally referred to as variable quantization. In variable quantization, the figures of merit in question are image quality (distortion) and output file size (rate). The problem is then to decrease image distortion for a target rate, or to decrease rate for a target distortion. Of particular interest is the latter, since it has direct application in minimizing bandwidth usage for images which are sent over computer networks. This also reduces the time to transmit the image, which is important when the network path includes slow speed links.

It is preferred that any technique for quantizing an image also be computationally efficient, especially when the quantization is performed on images which are generated dynamically, or images which cannot be stored in a caching

system. If the quantization is too slow, then any transmission time benefit realized from the reduction in rate is effectively annulled by the latency introduced in the quantization computation.

5

Accordingly, it is an object of the invention to provide a method for quantizing a JPEG image, which offers many of the benefits of variable quantization and is computationally efficient, while conforming to the widely used JPEG standard.

#### **SUMMARY OF THE INVENTION**

According to the present invention there is provided a method, which is directed towards regionally variable levels of compression. The method is directed to JPEG compression of an image frame divided up into  $8 \times 8$  pixel blocks  $B_{ij}$ , where  $i, j$  are integers covering all of the blocks in the image frame. The method includes forming a discrete cosine transform (DCT) of each block  $B_{ij}$  of the image frame to produce a matrix of blocks of transform coefficients  $D_{ij}$ . Next a visual importance,  $I_{ij}$ , is calculated for each block of the image, based upon assigning zeros for flat features and values approaching unity for sharply varying features. A global quantization matrix  $Q$  is then formed such that the magnitude of each quantization matrix coefficient  $Q_{ij}$  is inversely proportional to a visual importance  $I_{ij}$  of a corresponding DCT basis vector to the image. This local visual

importance is used during the quantization stage, where for regions of lower detail, more data is discarded, resulting in more aggressive compression. In the quantization stage the transform coefficients are quantized by dividing them by a factor  $S_{ij} Q$ , where  $S_{ij}$  is a linear scaling factor, to create a JPEG image file. This algorithm is unique in that it allows for the effect of variable-quantization to be achieved while still producing a file which conforms to the JPEG standard.

The visual importance,  $I_{ij}$ , may be determined by discrete edge detection and summation of transform coefficients. This determination of  $I_{ij}$  may include creating a  $24 \times 24$  matrix of image pixels of DCT coefficients centered on a block  $B_{ij}$ , where  $i$  and  $j = 1, 2, \dots, 8$ . The center  $10 \times 10$  matrix of the  $24 \times 24$  matrix may be convolved with an edge tracing kernel. The matrix values of the convolved matrix may be summed, and the summed value normalized to produce a visual importance,  $I_{ij}$ .

The quantization matrix,  $Q$ , may be formed by calculating an  $8 \times 8$  matrix  $A$  by calculating matrix elements  $A_{mn}$  according to the formula:

$$A_{mn} = \sum_{V(i,j)} I_{ij} (B_{ij})_{mn}.$$

Elements  $Q_{mn}$  of  $Q$  may then be calculated according to the formula:



$$Q_{mn} = \max(A_{mn}) / A_{mn}$$

and scaling values of  $Q_{mn}$  for all values of  $(m,n)$  except  $(0,0)$  in order to minimize the error between  $Q$  and a standard JPEG quantization matrix.

The linear scaling factor  $S_{ij}$  may be set equal to  $1_{ij}(S_{\max} - S_{\min}) + S_{\min}$ , where  $S_{\max}$  and  $S_{\min}$  are user selected.

Quantizing the blocks of DCT coefficients  $D_{ij}$  to produce quantized DCT coefficients  $T_{ijm}$ , where  $m$  and  $n$  refer to row and column, respectively, in each of the blocks may be accomplished by applying the formula.

$T_{ijmn} = \text{round}(D_{ijmn} / (S_{ij} * Q_{mn}))$ , where round denotes rounding to the nearest integer,

and if  $T_{ijmn} > 0$

calculate  $\text{round}(D_{ijmn} / (S_{ij} * Q_{mn}))$  and if equal to zero then set  $T_{ijmn} = 0$ , otherwise if

$$(\text{abs}(D_{ijmn}) - (2^{\lceil \lg(\text{abs}(D_{ijmn}) + 1) \rceil} - 1) - 1) \leq \text{abs}(D_{ijmn} - Q_{mn} * S_{ij} * \text{round}(D_{ijmn} / (S_{ij} * Q_{mn})))$$

then

$$T_{ijmn} = \text{sign}(D_{ijmn}) * (2^{\lceil \lg(\text{abs}(D_{ijmn}) + 1) \rceil} - 1) - 1.$$

According to another aspect of the invention there is provided a method of JPEG compression of a colour image represented by channels Y for greyscale data, and U and V each for colour, which comprises shrinking the colour channels U and V by a fraction of their size, forming a discrete cosine transform (DCT)  $D_{ij}$  for each block  $B_{ij}$  of each of channels Y, U and V and calculating a visual importance,  $I_{ij}$ , for each Y channel block of each image and setting  $I_{ij} = \max\{I_{ij} \text{ values for corresponding Y channel blocks}\}$  for blocks in the U and V channels. A global quantization matrix Q is formed for the Y channel block and one for channels U and V combined such that a magnitude of each quantization matrix coefficient  $Q_{ij}$  is inversely proportional to a visual importance  $I_{ij}$  to the image of a corresponding DCT basis vector. Next the transform coefficients for each of the Y, U and V channels are quantized by dividing them by a factor  $S_{ij} Q'$ , where  $S_{ij}$  is a linear scaling factor for each of channels Y, U and V and  $Q'$  is the quantization table for the associated channel being quantized. Finally, the quantized coefficients  $T_{ijmn}$  and  $Q' * S_{min}$  are entropy encoded, where  $S_{min}$  is a user selected minimum scaling factor for each of channels Y, U, and V, to create a JPEG image file for each of channels Y, U and V.

Preferably, the shrinking factor is 1/2.

In another aspect of the invention there is provided an apparatus for JPEG compression of an image frame divided up into a plurality of non-overlapping, tiled 8 x 8 pixel blocks  $B_{ij}$  where



## BRIEF DESCRIPTION OF THE DRAWINGS

Further features and advantages will be apparent from the following detailed description, given by way of example, of a preferred embodiment taken in conjunction with the accompanying  
5 drawings, wherein:

Fig. 1 is a schematic diagram showing a conventional JPEG system;

10 Fig. 2 is a flowchart showing the sequence of steps in the algorithm; and

15 Fig. 3 is a schematic diagram of the JPEG image compressor.

## DETAILED DESCRIPTION WITH REFERENCE TO THE DRAWINGS

An image frame is selected at step 40. The image frame is divided into non-overlapping tiled 8x8 pixel blocks  $B_{ij}$  at step 42 according to the JPEG standard.

20

For each 8x8 block  $B_{ij}$  in the image frame, a visual image importance  $I_{ij}$  is calculated at step 44. Note that the actual measure of visual importance is not important to the outline of the algorithm. The  $I_{ij}$  values exhaustively cover the  
25 range  $[0,1]$ , and are a measure of how aggressively the block can be quantized. A value of  $I_{ij} = 0$  indicates that the visual appearance of the block is rather insensitive to the level of

quantization, and a value of  $I_{ij} = 1$  indicates that the visual appearance of the block is very sensitive to the level of quantization.

5 One method of selecting the visual importance  $I_{ij}$  is based on a discrete edge-detection and summation technique. Consider a  $24 \times 24$  window  $W_{ij}$  on the image defined by the nine image blocks,  $B_{i-1,j-1}$ ,  $B_{i,j-1}$ ,  $B_{i+1,j-1}$ ,  $B_{i,j}$ ,  $B_{i+1,j}$ ,  $B_{i-1,j+1}$ ,  $B_{i,j+1}$ ,  $B_{i+1,j+1}$ . This window is centered around the block  $B_{ij}$ . The nine blocks are shown graphically in the following diagram:

$B_{i-1,j-1}$	$B_{i,j-1}$	$B_{i+1,j-1}$
$B_{i-1,j}$	$B_{i,j}$	$B_{i+1,j}$
$B_{i-1,j+1}$	$B_{i,j+1}$	$B_{i+1,j+1}$

From this  $24 \times 24$  window, a  $10 \times 10$  window  $V_{ij}$ , centered about  $B_{ij}$ , is then convolved with a standard Laplacian edge detection kernel  $G$ , to give  $H_{ij}$ . The edge detection kernel employed is,

$$G = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

25 and the convolution is given by,

$$H_{ij} = \sum_{m,n} V_{i-m,j-n} G_{ij}$$

This technique is essentially the discrete equivalent  
 5 of taking the second derivative of the image in both dimensions.  
 The output of the convolution  $H_{ij}$  is scaled to cover an 8-bit  
 range between 0 and 255, the same range taken by the actual  
 pixels in the image. The convolved values are then summed, and  
 the sum is divided by  $100 \times 255$  to scale the sum to the range 0 to  
 1. This scaled sum is denoted as  $K_{ij}$ . This sum is then  
 10 renormalized using the following function:

$$I_{ij} = \frac{K_{ij}(100 + C)}{100K_{ij} + C}$$

15 where  $C$  is equal to 14. This function is determined  
 statistically, and remaps the  $K_{ij}$  values such that they lie on a  
 normal distribution.

The above procedure is used to calculate  $I_{ij}$  for each  
 20 block in the image. The end result is a value for each  $I_{ij}$  which  
 is bounded on the region  $(0,1)$ , takes values of 0 for flat  
 blocks, and values approaching 1 for blocks that have lots of  
 sharp, short features (in other words have large second  
 derivatives).

The quantization matrix  $Q$  is determined at step 46. In one approach,  $Q$  is simply set equal to the standard JPEG quantization table, which is in general use, by the community. An example of a suitable such matrix is the following:

```

16,  11,  10,  16,  24,  40,  51,  61,
12,  12,  14,  19,  26,  58,  60,  55,
14,  13,  16,  24,  40,  57,  69,  56,
14,  17,  22,  29,  51,  87,  80,  62,
18,  22,  37,  56,  68, 109, 103,  77,
24,  35,  55,  64,  81, 104, 113,  92,
49,  64,  78,  87, 103, 121, 120, 101,
72,  92,  95,  98, 112, 100, 103,  99

```

In another approach, an image-specific quantization matrix is generated, where the magnitude of each quantization table coefficient is inversely proportional to the importance in the image of the corresponding basis vector.

One approach to generating an image-specific quantization matrix  $Q$  defines an 8x8 array such that each value  $A_{mn}$  is equal to the sum of the corresponding coefficients  $(m,n)$  in each block  $B_{ij}$ , weighted by the importance value  $I_{ij}$ :

$$A_{mn} = \sum_{\forall(i,j)} I_{ij} (B_{ij})_{mn}$$

After this summation, the matrix  $A$  holds relative counts of importance for each basis vector in the DCT transform.

This matrix is simply inverted and scaled entry-wise such that

$\bar{A}_{mn} = \max(A_{mn})/A_{mn}$ . In the cases where  $A_{mn}$  is zero,  $\bar{A}_{mn}$  is set to 255, which is the largest allowable value for an 8 bit number.

The values in  $\bar{A}_{mn}$  are then scaled such that the squared error between  $\bar{A}_{mn}$  and the standard JPEG quantization matrix is

5 minimized. The quantization matrix  $Q$  is then set equal to this scaled matrix. Note that this process is only performed on the AC coefficients, in other words for all values of  $(m,n)$  except  $(0,0)$ . For the  $(0,0)$  entry,  $Q_{00}$  is simply initialized to the corresponding value in the standard JPEG quantization table.

Each block  $B_{ij}$  is DCT transformed at step 48 according to the JPEG standard to produce DCT coefficients  $D_{ij}$ .

For each block  $B_{ij}$  in the image, a value  $S_{ij}$  is calculated at step 50 where  $S_{ij} = I_{ij} * (S_{\max} - S_{\min}) + S_{\min}$ . The parameters  $S_{\max}$  and  $S_{\min}$  are user specified and in effect define the bounds over which the image will be variably quantized. This method is preferably used to remove redundant data from an existing compressed JPEG by letting  $S_{\min}$  be equal to the actual scaling value used in compressing the image originally, and using a user-defined value for  $S_{\max}$ .

Each block  $B_{ij}$  in the image is "pseudo-quantized" at step 56 with the quantization table  $Q_{mn} * S_{ij}$ . This pseudo-quantization in effect emulates variable quantization. If one lets  $D_{ij}$  be the original unquantized DCT transformed image block,



and  $T_{ij}$  the quantized DCT transformed block at step 54, then the algorithm for the pseudo-quantization can be described as given next.

5           The algorithm has three distinct quantization steps. In the first step, the coefficients in the block  $B_{ij}$  are quantized using the standard JPEG quantization function with  $S_{min}$  as the scaling value:

10       for each block  $D_{ij}$  do

          for each coefficient  $D_{ijmn}$  in block  $D_{ij}$  do

$$T_{ijmn} = \text{round}(D_{ijmn} / (Q_{mn} * S_{min}))$$

where round denotes rounding to the nearest integer.

15       In the next step, if any coefficient  $T_{ijmn}$  is  $> 0$ , then

$$\text{if } \text{round}(D_{ijmn} / (Q_{mn} * S_{ij})) = 0 \text{ then } T_{ijmn} = 0$$

20       In the third and final step, if  $T_{ij}$  is still greater than zero, and if the coefficient can be rounded down by one logarithm base-2 and not exceed the rounding error introduced by the quantization with the local quantization table, then it is so rounded down:

$$\begin{aligned} &\text{if } (\text{abs}(D_{ijmn}) - (2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn}))+1))-1}-1)) \\ &\leq \text{abs}(D_{ijmn} - Q_{mn} * S_{ij} * \text{round}(D_{ijmn} / (Q_{mn} * S_{ij}))) \end{aligned}$$

then

$$T_{ij} = \text{sign}(D_{ijmn}) * (2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))-1)-1})$$

In the above calculations,  $Q_{mn} * \text{round}(D_{ijmn} / (S_{ij} * Q_{mn}))$  is the  
5 reconstructed coefficient after quantization by the local  
quantization table, and

$$\text{abs}(D_{ijmn} - Q_{mn} * S_{ij} * \text{round}(D_{ijmn} / (S_{ij} * Q_{mn})))$$

10 is the absolute error introduced by quantization. Furthermore,

$$\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))$$

15 is the logarithm base-2 of the magnitude of the coefficient, and,

$$(2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))-1})-1)$$

is the magnitude of the coefficient rounded down by a logarithm  
base-2.

20 Thus,

$$\text{abs}(D_{ijmn}) - (2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))-1})-1)$$

25 is the absolute error introduced by rounding down by one  
logarithm base-2.

The algorithm in its entirety is:

```

for each block  $D_{ij}$  do
{
5   for each coefficient  $D_{ijmn}$  in block  $D_{ij}$  do
      {
           $T_{ijmn} = \text{round}(D_{ijmn} / (S_{mn} * Q_{mn}))$ 
          if  $T_{ijmn} > 0$  then
              {
                  if  $\text{round}(D_{ijmn} / (S_{ij} * Q_{mn})) = 0$  then  $T_{ijmn} = 0$ 
                  else
                      {
                          if  $(\text{abs}(D_{ijmn}) - (2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))-1)-1})) \leq$ 
                           $\text{abs}(D_{ijmn} - Q_m * S_{ij} * \text{round}(D_{ijmn} / (S_{ij} * Q_{mn})))$ 
                          then
                               $T_{ijmn} = \text{sign}(D_{ijmn}) * (2^{(\text{ceil}(\lg(\text{abs}(D_{ijmn})+1))-1)-1})$ 
                          }
                      }
              }
      }
}
20 }
```

The above pseudo-code has the effect of zeroing any coefficients that would have been zeroed if  $D_{ijmn}$  were quantized with  $Q_{mn} * S_{ij}$ , but were not zeroed when quantized with  $Q_{mn} * S_{mn}$ . Also, it rounds down in magnitude (by one power of two) any coefficient that may be so rounded and not introduce more relative error in reconstruction than if that coefficient were truly quantized by

$Q_{mn} * S_{ij}$ . This has the net effect of pseudo-quantizing  $D_{ij}$  with  $Q * S_{ij}$ , while actually quantizing the coefficients with  $Q * S_{min}$ .

Finally, the quantized blocks  $T_{ijmn}$  and the global quantization table  $Q * S_{min}$  are entropy encoded at step 58 to create a JPEG image file 60 in accordance with the JPEG algorithm while still producing a fully compliant JFIF stream.

It should be noted that the algorithm is particularly useful in optimizing JPEG images that have already been quantized using the standard JPEG quantization table at a level  $S_{min}$ . By definition  $S_{ij} \geq S_{min}$ , hence the algorithm guarantees that the optimized JPEG will never be larger in size than the original JPEG, and will in almost all instances be smaller. At the same time the pseudo-quantization ensures that the image quality remains essentially unchanged to the human observer.

For the sake of clarity, the algorithm has been presented assuming the image contains a single 8-bit channel per pixel, in other words it is a greyscale image. However, the algorithm is easily extended to full color (3 channel) images, and more generally,  $n$  channel images with few adjustments to the process. In general, the algorithm is simply applied to each channel independently, where the visual importance values are calculated on the luminance channel. A single quantization matrix  $Q$  can be employed for all channels, or alternatively, a separate quantization matrix can be used for each channel. Likewise,  $S_{min}$

and  $S_{\max}$  can either be the same for all channels, or vary from channel to channel.

It is common practice to sub-sample one or more channels when color images are coded. The algorithm can still be employed in this case. An example using a full color, 3-channel image will be described.

A common color scheme to represent a color image is known as YUV. Here, Y stands for the luminance channel (or the greyscale data), and U and V are the blue and red chrominance (color) channels respectively. Since the human visual system perceives luminance information much better than color data, the U and V channels are typically sub-sampled by an integer factor, normally 2, to improve compression. In this case, in the original pixel domain the image is shrunk to half its original size, and then DCT transformed. When decoding, the inverse transform is applied and the plane is expanded by twice its size before merging the three channels to reconstruct the original image.

Because of the subsampling, there may be up to four Y channel blocks that correspond to the same region of an image covered by one U and V block. In this case, the visual importance  $I_{ij}$  that is used is simply given as,

$\max \{\text{all corresponding } I_{ij} \text{ values from the Y channel}\}.$

Referring to Figure 3 the apparatus for JPEG

Compression using the above algorithm consists of a frame grabber  
80 into which non-overlapping, tiled, 8 x 8 image pixel blocks  $B_{ij}$   
are stored temporarily. Each block,  $B_{ij}$ , the digital cosine  
5 transform (DCT) is calculated by DCT transformer 82 and the  
resultant transform coefficients  $D_{ijmn}$  stored in memory 84. A  
visual importance calculator 86 calculates values of the visual  
importance,  $I_{ij}$ , for each block  $B_{ij}$ . A global quantization  
calculator 87 calculates elements  $Q_{ij}$  of a global quantization  
10 matrix utilizing,  $I_{ij}$ , and  $B_{ij}$ . A linear scaling factor calculator  
89 uses user set values of  $S_{ijmin}$  and  $S_{ijmax}$  set in blocks 124 and  
126, respectively, and  $I_{ij}$  to determine  $S_{ij}$  in calculator 128 for  
quantized blocks  $T_{ij}$ .

More particularly, values of the quantization matrix  $Q_{ij}$   
are calculated by first forming the sum of the product of the  
visual importance  $I_{ij}$  and the elements of  $B_{ij}$  in block 88 to form  
the elements  $A_{mn}$  in an 8 x 8 array which are stored in memory 100.  
The maximum value "Max  $A_{mn}$ " in the array is selected by Max  $A_{mn}$   
15 selector 102. The elements  $Q_{mn}$  of the quantization matrix  $Q$  are  
calculated as  $(\text{Max } A_{mn})/A_{mn}$  in block 104.

In block 106, the quotient of  $D_{ijmn}/(S_{ijmin} * Q_{mn})$  is rounded  
to the nearest integer yielding elements  $T_{ijmn}$ . In comparator 108,  
25 the calculated value of  $T_{ijmn}$  is compared with zero and, if greater  
than zero, in block 110 the quotient  $D_{ijmn}/(S_{ij} * Q_{mn})$  is calculated and

then rounded to the nearest integer. If the quotient  $D_{ijmn}/(S_{ij} * Q_{mn})$  equals zero, then  $T_{ijmn}$  is set equal to zero at block 112. If the quotient  $D_{ijmn}/(S_{ij} * Q_{mn})$  is not equal to zero, at block 110, then the value of the rounded value of the latter quotient is transferred to block 116. Values calculated in blocks 116 and 118 are compared in calculator 120 and if the value calculated in block 116 is less than or equal to the value calculated in block 118, then the value of  $T_{ijmn}$  is set equal to  $\text{sign}(D_{ijmn}) * (2^{(\text{ceil}(\lg(\text{abs}D_{ijmn})+1)-1)} - 1)$ . The blocks of quantized coefficients  $T_{ij}$  and the global quantization table  $Q * S_{\min}$  are entropy encoded by entropy encoder 113.

Accordingly, while this invention has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense. Various modifications of the illustrative embodiments, as well as other embodiments of the invention, will be apparent to persons skilled in the art upon reference to this description. It is therefore contemplated that the appended claims will cover any such modifications or embodiments as they fall within the true scope of the invention.